Reflection positivity, rank connectivity, and homomorphism of graphs *

MICHAEL FREEDMAN, LÁSZLÓ LOVÁSZ
Microsoft Research
One Microsoft Way
Redmond, WA 98052
and
ALEXANDER SCHRIJVER
CWI Amsterdam
The Netherlands

March 2004

MSR-TR-2004-41

Abstract

It is shown that a graph parameter can be realized as the number of homomorphisms into a fixed (weighted) graph if and only if it satisfies two linear algebraic conditions: reflection positivity and exponential rank-connectivity. In terms of statistical physics, this can be viewed as a characterization of partition functions of vertex models.

1 Introduction

For two finite graphs G and H, let $\hom(G, H)$ denote the number of homomorphisms (adjacency-preserving mappings) from G to H. Many interesting graph parameters can be expressed in terms of the number of homomorphisms into a fixed graph: for example, the number of colorings with a number of colors is the number of homomorphisms into a complete graph. Further examples for the occurrence of these numbers in graph theory will be discussed in Section 3.

---

*AMS Subject Classification: Primary 05C99, Secondary 82B99

Another source of important examples is statistical physics, where partition functions of various models can be expressed as graph homomorphism functions. For example, let G be a $n \times n$ grid, and suppose that every node of G (every "site") can be in one of two states, "UP" or "DOWN". The properties of the system are such that no two adjacent sites can be "UP". A "configuration" is a valid assignment of states to each node. The number of configurations is the number of independent sets of nodes in G, which in turn can be expressed as the number of homomorphisms of G into the graph H consisting of two nodes, "UP" and "DOWN", connected by an edge, and with an additional loop at "DOWN". To capture more interesting physical models, so called "vertex models", one needs to extend the notion of graph homomorphism to the case when the nodes and edges of H have weights (see Section 2.1).

Which graph parameters can be represented as homomorphism functions into weighted graphs? This question is motivated, among others, by the problem of physical realizability of certain graph parameters. Two necessary conditions are easy to prove:

(a) The interaction between two parts of a graph separated by k nodes is bounded by a simple exponential function of k (this will be formalized as *rank-connectivity* in Section 2.2).

(b) Another necessary condition, which comes from statistical mechanics, as well as from extremal graph theory, is *reflection positivity*. Informally, this means that if a system has a 2-fold symmetry, then its partition function is positive. We'll formulate a version of this in section 2.2.

The main result of this paper is to prove that these two necessary conditions are also sufficient. The proof makes use of a simple kind of (commutative, finite dimensional) C^*-algebras.

2 Homomorphisms, rank-connectivity and reflection positivity

2.1 Weighted graph homomorphisms

A *weighted graph* H is a graph with a positive real weight $\alpha_H(i)$ associated with each node i and a real weight $\beta_H(i,j)$ associated with each edge ij.

Let G be an unweighted graph (possibly with multiple edges, but no loops) and H, a weighted graph. To every homomorphism $\phi : V(G) \rightarrow$

$V(H)$, we assign the weight

$$w(\phi) = \prod_{u \in V(G)} \alpha_H(\phi(u)) \prod_{uv \in E(G)} \beta_H(\phi(u), \phi(v)), \qquad (1)$$

and define

$$\hom(G, H) = \sum_{\substack{\phi:\ V(G) \to V(H) \\ \text{homomorphism}}} w(\phi). \qquad (2)$$

If all the node-weights and edge-weights in H are 1, then this is the number of homomorphisms from G into H (with no weights).

For the purpose of this paper, it will be convenient to assume that H is a complete graph with a loop at all nodes (the missing edges can be added with weight 0). Then the weighted graph H is completely described by a positive integer $d = |V(H)|$, the positive real vector $a = (\alpha_1, \ldots, \alpha_d) \in \mathbf{R}^d$ and the real symmetric matrix $B = (\beta_{ij}) \in \mathbf{R}^{d \times d}$. The graph parameter $\hom(., H)$ will be denoted by f_H or $f_{B,a}$.

2.2 Connection matrices of a graph parameter

A *graph parameter* is a function on finite graphs (invariant under graph isomorphism). We allow multiple edges in our graphs, but no loops. A graph parameter f is called *multiplicative*, if for the disjoint union $G_1 \cup G_2$ of two graphs G_1 and G_2, we have $f(G_1 \cup G_2) = f(G_1)f(G_2)$.

A *k-labeled graph* ($k \geq 0$) is a finite graph in which k nodes are labeled by $1, 2, \ldots k$ (the graph can have any number of unlabeled nodes). Two k-labeled graphs are *isomorphic*, if there is a label-preserving isomorphism between them. We denote by K_k the k-labeled complete graph on k nodes, and by O_k, the k-labeled graph on k-nodes with no edges. In particular, $K_0 = O_0$ is the graph with no nodes and no edges.

Let G_1 and G_2 be two k-labeled graphs. Their *product* G_1G_2 is defined as follows: we take their disjoint union, and then identify nodes with the same label. Hence for two 0-labeled graphs, $G_1G_2 = G_1 \cup G_2$ (disjoint union).

Now we come to the construction that is crucial for our treatment. Let f be any graph parameter. For every integer $k \geq 0$, we define the following (infinite) matrix $M(f, k)$. The rows and columns are indexed by isomorphism types of k-labeled graphs. The entry in the intersection of the row corresponding to G_1 and the column corresponding to G_2 is $f(G_1G_2)$. We call the matrices $M(f, k)$ the *connection matrices* of the graph parameter f.

We will be concerned with two properties of connection matrices, namely their rank and positive semidefiniteness. The rank $r_f(k) = \text{rk}(M(f, k))$, as a

3

function of k, will be called the *rank connectivity function* of the parameter f. This may be infinite, but for many interesting parameters it is finite, and its growth rate will be important for us. We say that a set of k-labeled graphs *spans* if the corresponding rows of $M(f, k)$ have rank $r_f(k)$.

We say that a graph parameter f is *reflection positive*, if $M(f, k)$ is positive semidefinite for every k. This property is closely related to the reflection positivity property of certain statistical physical models. Indeed, it implies that for any k-labeled graph G, $f(GG) \geq 0$ (since $f(GG)$ is a diagonal entry of a positive semidefinite matrix). Here the second copy of G can be thought of as a "reflection" (in the set of labeled nodes) of the first.

The condition that $f(GG) \geq 0$ is weaker than the condition that $M(f, k)$ is positive semidefinite, but we can strengthen it as follows. Let $x = (x_G)$ be a complex vector indexed by k-labeled graphs with $\sum_G |x_G|^2 = 1$. The formal linear combination $X = \sum_G x_G G$ can be thought of as a "quantum k-labeled graph", and then $X\overline{X} = \sum_{G_1, G_2} x_{G_1} \overline{x}_{G_2}(G_1 G_2)$ is a quantum graph that can be obtained by gluing together X and its reflection. (Note that we took complex conjugation of the coefficient to build the "mirror copy" of X.) Now if we extend f linearly over quantum graphs, then

$$f(XX) = x^{\mathsf{T}} M(f, k)\overline{x},$$

showing that the non-negativity of f over symmetric quantum graphs is equivalent to reflection positivity as we defined it.

2.3 Simple properties of connection matrices

We give a couple of simple facts about connection matrices of a general graph parameter.

Proposition 2.1 *Let f be a graph parameter that is not identically 0. Then f is multiplicative if and only if $M(f, 0)$ is positive semidefinite, $f(K_0) = 1$, and $r_f(0) = 1$.*

Proof. If f is multiplicative, then $f(K_0)^2 = f(K_0)$, showing that $f(K_0) \in \{0, 1\}$. If $f(K_0) = 0$, then the relation $f(G) = f(GK_0) = f(G)f(K_0)$ implies that $f(G) = 0$ for every G, which was excluded. So $f(K_0) = 1$. Furthermore, $f(G_1 G_2) = f(G_1)f(G_2)$ for any two 0-labeled graphs G_1 and G_2, which implies that $M(f, 0)$ has rank 1 and is positive semidefinite.

Conversely, since $M(f, 0)$ is symmetric, the assumption that $r_f(0) = 1$ implies that there is a graph parameter ϕ and a constant c such that $f(G_1 G_2) = c\phi(G_1)\phi(G_2)$. Since $M(f, 0)$ is positive semidefinite, we have

$c > 0$, and so we can normalize ϕ to make $c = 1$. Then $f(K_0) = f(K_0 K_0) = \phi(K_0)^2$, whence $\phi(K_0) \in \{-1, 1\}$. We can replace ϕ by $-\phi$, so we may assume that $\phi(K_0) = 1$. Then $f(G) = f(GK_0) = \phi(G)\phi(K_0) = \phi(G)$ for every G, which shows that f is multiplicative. $\qquad \square$

Proposition 2.2 *Let f be a multiplicative graph parameter and $k, l \geq 0$. Then*

$$r_f(k + l) \geq r_f(k) \cdot r_f(l).$$

We'll see (see Claim 4.7 below) that in the case when f is reflection positive, the sequence $r_f(k)$ logconvex. We don't know if this property holds in general.

Proof. Let us call a $(k + l)$-labeled graph *separated*, if every component of it contains either only nodes with label at most k, or only nodes with label larger than k. Consider the submatrix of $M(f, k+l)$ formed by the separated rows and columns. By multiplicativity, this submatrix is the Kronecker (tensor) product of $M(f, k)$ and $M(f, l)$, so its rank is $r_f(k) \cdot r_f(l)$. $\qquad \square$

2.4 Connection matrices of homomorphisms

Fix a weighted graph $H = (a, B)$. For any k-labeled graph G and mapping $\phi : [1, k] \to V(H)$, let

$$\hom_\phi(G, H) = \sum_{\substack{\psi:\ V(G) \to V(H) \\ \psi \text{ extends } \phi}} w(\psi). \tag{3}$$

So $\hom(G, H) = \sum_\phi \hom_\phi(G, H)$.

Theorem 2.3 *The graph parameter f_H is reflection positive and $r_{f_H}(k) \leq |V(H)|^k$.*

Proof. For any two k-labeled graph G_1 and G_2 and $\phi : [1, k] \to V(H)$,

$$\hom_\phi(G_1 G_2, H) = \frac{1}{\prod_{i=1}^k \alpha_H(\phi(i))} \hom_\phi(G_1, H) \hom_\phi(G_2, H). \tag{4}$$

The decomposition (3) writes the matrix $M(f, k)$ as the sum of $|V(H)|^k$ matrices, one for each mapping $\phi : [1, k] \to V(H)$; (4) shows that these matrices are positive semidefinite and have rank 1. $\qquad \square$

The main result of this paper is a converse to Theorem 2.3:

Theorem 2.4 *Let f be a reflection positive graph parameter for which there exists a positive integer q such that $r_f(k) \leq q^k$ for every $k \geq 0$. Then there exists a weighted graph H such that $f = f_H$.*

3 Examples

We start with an example showing that exponential growth of rank connectivity is not sufficient in itself to guarantee that a graph parameter is a homomorphism function.

Example 3.1 (Matchings) Let $\Phi(G)$ denote the number of perfect matchings in the graph G. It is trivial that $\Phi(G)$ is multiplicative. We claim that its node-rank-connectivity is exponentially bounded:

$$r_\Phi(k) = 2^k.$$

Let G be a k-labeled graph, let $X \subseteq [1, k]$, and let $\Phi(G, X)$ denote the number of matchings in G that match all the unlabeled nodes and the nodes with label in X, but not any of the other labeled nodes. Then we have for any two k-labeled graphs G_1, G_2

$$\Phi(G_1 G_2) = \sum_{X_1 \cap X_2 = \emptyset,\, X_1 \cup X_2 = [1,k]} \Phi(G_1, X_1)\Phi(G_2, X_2).$$

This can be read as follows: The matrix $M(\Phi, k)$ can be written as a product $N^\mathsf{T} W N$, where N has infinitely many rows indexed by k-labeled graphs, but only 2^k columns, indexed by subsets of $[1, k]$,

$$N_{G,X} = \Phi(G, X),$$

and W is a symmetric $2^k \times 2^k$ matrix, where

$$W_{X_1, X_2} = \begin{cases} 1 & \text{if } X_1 = [1, k] \setminus X_2, \\ 0 & \text{otherwise.} \end{cases}$$

Hence the rank of $M(\Phi, k)$ is at most 2^k (it is not hard to see that in fact equality holds).

On the other hand, let us consider K_1 and K_2 as 1-labeled graphs. Then the submatrix of $M(\Phi, 1)$ indexed by K_1 and K_2 is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

which is not positive semidefinite. Thus $\Phi(G)$ cannot be represented as a homomorphism function.

Example 3.2 (Flows) The following example shows that there are important graph parameters that are not defined as homomorphism functions, but that can also be represented as homomorphism functions in a nontrivial way. In fact, it is easier to check that the conditions in Theorem 2.4 hold than to show that these graph parameters are homomorphism functions, and they first came up as counterexample candidates.

Let us start with a simple special case. Let $f(G) = 1$ if G is eulerian (i.e., all nodes have even degree), and $f(G) = 0$ otherwise. To represent this function as a homomorphism function, let

$$a = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

It is not hard to see that for the weighted graph $H = (a, B)$ we have $\hom(G, H) = f(G)$.

It follows that for this function, reflection positivity holds, and the rank-connectivity is at most 2^k.

This example can be generalized quite a bit. Let Γ be a finite abelian group and let $S \subseteq \Gamma$ be such that S is closed under inversion. For any graph G, fix an orientation of the edges. An S-*flow* is an assignment of an element of S to each edge such that for each node v, the product of elements assigned to edges entering v is the same as the product of elements assigned to the edges leaving v. Let $f(G)$ be the number of S-flows. This number is independent of the orientation.

The choice $\Gamma = \mathbf{Z}_2$ and $S = \mathbf{Z}_2 \setminus \{0\}$ gives the special case above (incidence function of eulerian graphs). If $\Gamma = S = \mathbf{Z}_2$, then $f(G)$ is the number of eulerian subgraphs of G. Perhaps the most interesting special case is when $|\Gamma| = t$ and $S = \Gamma \setminus \{0\}$, which gives the number of nowhere zero t-flows.

Surprisingly, this parameter can be described as a homomorphism function. Let Γ^* be the character group of Γ. Let H be the complete directed graph (with all loops) on Γ^*. Let $\alpha_\chi := 1/|\Gamma|$ for each $\chi \in \Gamma^*$, and let

$$\beta_{\chi,\chi'} := \sum_{s \in S} \chi^{-1}(s)\chi'(s),$$

for $\chi, \chi' \in \Gamma^*$.

We show that

$$f = \hom(., H). \tag{5}$$

Let $n = |V(G)|$ and $m = |\Gamma|$. For any coloring $\phi : V(G) \to S$ and node $v \in V(G)$, let

$$\partial_\phi(v) = \sum_{\substack{u \in V(G) \\ uv \in E(G)}} \phi(uv) - \sum_{\substack{u \in V(G) \\ vu \in E(G)}} \phi(vu).$$

7

So ϕ is an S-flow if an only if $\partial_\phi = 0$. Consider the expression

$$A = \sum_{\phi:\ E(G) \to S} \prod_{v \in V(G)} \sum_{\chi \in \Gamma^*} \chi(\partial_\phi(v)).$$

The summation over χ is 0 unless $\partial_\phi(v) = 0$, in which case it is m. So the product over $v \in V(G)$ is 0 unless ϕ is an S-flow, in which case it is m^n. So $A \cdot m^{-n}$ counts S-flows.

On the other hand, we can expand the product over $v \in V(G)$; each term will correspond to a choice of a character ψ_v for each v, and so we get

$$A = \sum_{\phi:\ E(G) \to S} \sum_{\psi:\ V(G) \to \Gamma^*} \prod_{v \in V(G)} \psi_v(\partial_\phi(v)).$$

Here (using that ψ_v is a character)

$$\psi_v(\partial_\phi(v)) = \prod_{\substack{u \in V(G) \\ uv \in E(G)}} \psi_v(\phi(uv)) \prod_{\substack{u \in V(G) \\ vu \in E(G)}} \psi_v(\phi(vu))^{-1},$$

so we get that

$$A = \sum_{\phi:\ E(G) \to S} \sum_{\psi:\ V(G) \to \Gamma^*} \prod_{uv \in E(G)} \psi_v(\phi(uv))\psi_u(\phi(uv))^{-1}.$$

Interchanging the summation, the inner sum factors:

$$\sum_{\phi:\ E(G) \to S} \prod_{uv \in E(G)} \psi_v(\phi(uv))\psi_u(\phi(uv))^{-1} = \prod_{uv \in E(G)} \sum_{s \in S} \psi_v(s)\psi_u(s)^{-1}$$

$$= \prod_{uv \in E(G)} \beta_{\psi_u \psi_v} = m^n w(\psi),$$

showing that

$$A = m^n \text{hom}(G, H).$$

This proves (5).

Example 3.3 (The role of multiple edges) Let us give an example of a reflection positive graph parameter f for which $r_f(k)$ is finite for every k but has superexponential growth. The example also shows that we have to be careful with multiple edges. Let, for each graph G, G' denote the graph which we obtain from G by keeping only one copy of each parallel class of edges. Let

$$f(G) = 2^{-|E(G')|}.$$

8

It is not hard to see that the connection matrices $M(f, k)$ are positive semidefinite. This graph parameter is, in fact, the limit of parameters of the form $\hom(., H)$: take homomomorphisms into a random graph $H = G(n, 1/2)$, with all node weights $= 1/n$ and all edge-weights 1. Furthermore, it is not hard to see that the rank of $M(f, k)$ is $2^{\binom{k}{2}}$. This is finite but superexponential, so the parameter is not of the form $\hom(., H)$.

Note, however, that for a simple graph G (*i.e.*, if G has no multiple edges), $f(G) = 2^{|E(G)|}$ can be represented as the number of homomorphisms into the graph consisting of a single node with a loop, where the node has weight 1 and the loop has weight $1/2$.

Example 3.4 (Chromatic polynomial) The following example is from [1]. Let $p(G) = p(G, x)$ denote the chromatic polynomial of the graph G. For every fixed x, this is a multiplicative graph parameter. To describe its rank-connectivity, we need the following notation. For $k, q \in \mathbf{Z}_+$, let B_{kq} denote the number of partitions of a k-element set into at most q parts. So $B_k = B_{kk}$ is the k-th Bell number. With this notation,

$$r_p(k) = \begin{cases} B_{kx} & \text{if } x \text{ is a nonnegative integer,} \\ B_k & \text{otherwise.} \end{cases}$$

Note that this is always finite, but if $x \notin \mathbf{Z}_+$, then it grows faster than c^k for every c.

Furthermore, $M(p, k)$ is positive semidefinite if and only if either x is a positive integer or $k \leq x + 1$. The parameter $M(p, k)$ is reflection positive if and only if this holds for every k, i.e., if and only if x is a nonnegative integer, in which case indeed $p(G, x) = \hom(G, K_x)$.

Example 3.5 (Homomorphisms into infinite graphs) We can extend the definition of $\hom(G, H)$ to infinite weighted graphs H provided the node and edge-weights form sufficiently fast convergent sequences. Reflection positivity remains valid, but the rank of $M(f_H, k)$ will become infinite, so this graph parameter cannot be represented by a finite H.

More generally, let $a > 0$, $I = [0, a]$, and let $W : I \times I \to \mathbf{R}$ be a measurable function such that for every $n \in \mathbf{Z}_+$,

$$\int_0^a \int_0^a |W(x, y)|^n \, dx \, dy < \infty.$$

Then we can define a graph parameter f_W as follows. Let G be a finite

9

graph on n nodes, then

$$f_W(G) = \int_{I^n} \prod_{ij \in E(G)} W(x_i, x_j)\, dx_1 \ldots dx_n.$$

It is easy to see that for every weighted (finite or infinite) graph H, the graph parameter f_H is a special case. Furthermore, f_W is reflection positive. However, it can be shown that the graph parameter in Example 3.3 cannot be represented in this form [2].

4 Proof of Theorem 2.4

4.1 The algebra of graphs

In this first part of the proof, we only assume that for every k, $M(f, k)$ is positive semidefinite, has finite rank r_k, and $r_0 = 1$. We know that f is multiplicative, and hence $f(K_0) = 1$. We can replace the parameter $f(G)$ by $f(G)/f(K_1)^{-|V(G)|}$; this can be reversed by scaling of the node-weights of the target graph H representing f, once we have it constructed. So we may assume that $f(K_1) = 1$. Combined with multiplicativity, this implies that we can delete (or add) isolated nodes from any graph G without changing $f(G)$.

It will be convenient to put all k-labeled graphs into a single structure as follows. By a *partially labeled graph* we mean a finite graph in which some of the nodes are labeled by distinct nonnegative integers. Two partially labeled graphs are *isomorphic* if there is an isomorphism between them preserving all labels. For two partially labeled graphs G_1 and G_2, let $G_1 G_2$ denote the partially labeled graph obtained by taking the disjoint union of G_1 and G_2, and identifying nodes with the same label. For every finite set $S \subseteq \mathbf{Z}_+$, we call a partially labeled graph S-*labeled*, if its labels form the set S.

Let $\mathcal{G}$ denote the (infinite dimensional) vector space of formal linear combinations (with real coefficients) of partially labeled graphs. We can turn $\mathcal{G}$ into an algebra by using $G_1 G_2$ introduced above as the product of two generators, and then extending this multiplication to the other elements linearly. Clearly $\mathcal{G}$ is associative and commutative, and the empty graph is a unit element.

For every finite set $S \subseteq \mathbf{Z}_+$, the set of all formal linear combinations of S-labeled graphs forms a subalgebra $\mathcal{G}(S)$ of $\mathcal{G}$. The graph with $|S|$ nodes labeled by S and no edges is a unit in this algebra, which we denote by U_S.

We can extend f to a linear functional on $\mathcal{G}$, and define an inner product

$$\langle x, y \rangle = f(xy)$$

for $x, y \in \mathcal{G}$. By our hypothesis that f is reflection positive, this inner product is positive semidefinite, i.e., $\langle x, x \rangle \geq 0$ for all $x \in \mathcal{G}$. Indeed, if $x = \sum_{G \in \mathcal{G}} x_G G \in \mathcal{G}$ (with a finite number of nonzero terms), then

$$\langle x, x \rangle = f(xx) = \sum_{G_1, G_2 \in \mathcal{G}} x_{G_1} x_{G_2} f(G_1 G_2) \geq 0,$$

since the quadratic form is of the form $x^{\mathsf{T}} M(f, k) x$ for a large enough k, and $M(f, k)$ is positive semidefinite.

Let

$$\mathcal{K} = \{x \in \mathcal{G} : \ f(xy) = 0 \ \forall y \in \mathcal{G}\}$$

be the annihilator of $\mathcal{G}$. It follows from the positive semidefiniteness of the inner product that $x \in \mathcal{K}$ could also be characterized by $f(xx) = 0$.

Clearly, $\mathcal{K}$ is an ideal in $\mathcal{G}$, so we can form the quotient algebra $\hat{\mathcal{G}} = \mathcal{G}/\mathcal{K}$. We can also define $\hat{\mathcal{G}}(S) = \mathcal{G}(S)/\mathcal{K}$. It is easy to check that every graph U_S has the same image under this factorization, namely unit element u of $\hat{\mathcal{G}}$. Furthermore, if $x \in \mathcal{K}$ then $f(x) = f(xu) = 0$, and so f can also be considered as a linear functional on $\hat{\mathcal{G}}$. We denote by $\hat{G}$ the element of $\hat{\mathcal{G}}$ corresponding to the partially labeled graph G.

Next, note that $\hat{\mathcal{G}}(S)$ is a finite dimensional commutative algebra of dimension $r_{|S|}$, with the positive definite inner product $\langle x, y \rangle = f(xy)$.

Let $S \subseteq T$ be finite subsets of $\mathbf{Z}_+$. Then $\hat{\mathcal{G}}(S) \subseteq \hat{\mathcal{G}}(T)$. Indeed, every S-labeled graph G can be turned into a T-labeled graph G' by adding $|T \setminus S|$ new isolated nodes, and label them by the elements of $T \setminus S$. It is straightforward to check that $G - G' \in \mathcal{K}$, and so G and G' correspond to the same element of $\hat{\mathcal{G}}$.

We'll also need the orthogonal projection π_S of $\hat{\mathcal{G}}$ to the subalgebra $\hat{\mathcal{G}}(S)$. This has a very simple combinatorial description. For every partially labeled graph G and $S \subseteq \mathbf{Z}_+$, let G_S denote the S-labeled graph obtained by deleting the labels not in S; then $\pi_S(\hat{G}) = \hat{G}_S$.

To see this, let G be any partially labeled graph. Then

$$\langle G_S, G - G_S \rangle = f(G_S(G - G_S)) = f(G_S G) - f(G_S G_S) = 0,$$

since GG_S and $G_S G_S$ are isomorphic graphs. Hence

$$\langle \hat{G}_S, \hat{G} - \hat{G}_S \rangle = 0,$$

11

showing that $\hat{G}_S$ is indeed the orthogonal projection of $\hat{G}$ onto $\hat{\mathcal{G}}(S)$.

We'll be interested in the idempotent elements of $\hat{\mathcal{G}}$. If p is idempotent, then

$$f(p) = f(pp) = \langle p, p \rangle > 0.$$

For two idempotents p and q, we say that q *resolves* p, if $pq = q$. It is clear that this relation is transitive.

Let S be a finite subset of $\mathbf{Z}_+$, and set $r = r_{|S|}$. Since the algebra $\hat{\mathcal{G}}(S)$ is finite dimensional and commutative, and all its elements are self-adjoint with respect to the positive definite inner product $\langle ., . \rangle$, it has a (uniquely determined) basis $\mathcal{P}_S = \{p_1^S, \ldots, p_r^S\}$ such that $(p_i^S)^2 = p_i^S$ and $p_i^S p_j^S = 0$ for $i \neq j$. We denote by $\mathcal{P}_{T,p}$ the set of all idempotents in $\mathcal{P}_T$ that resolve a given idempotent p. If $|T| = |S| + 1$, then the number of elements in $\mathcal{P}_{T,p}$ will be called the *degree* of $p \in \mathcal{P}_S$, and denoted by $\deg(p)$. Obviously this value is independent of which $(|S| + 1)$-element superset T of S we are considering.

Claim 4.1 *Let r be any idempotent element of $\mathcal{G}(S)$. Then r is the sum of those idempotents in $\mathcal{P}_S$ that resolve it.*

Indeed, we can write

$$r = \sum_{p \in \mathcal{P}_S} \mu_p p$$

with some scalars μ_p. Now using that r is idempotent:

$$r = r^2 = \sum_{p,p' \in \mathcal{P}_S} \mu_p \mu_{p'} pp' = \sum_{p \in \mathcal{P}_S} \mu_p^2 p,$$

which shows that $\mu_p^2 = \mu_p$ for every p, and so $\mu_p \in \{0, 1\}$. So r is the sum of some subset $X \subseteq \mathcal{P}_S$. It is clear that $rp = p$ for $p \in X$ and $rp = 0$ for $p \in \mathcal{P}_S \setminus X$, so X consists of exactly those elements of $\mathcal{P}_S$ that resolve q.

As a special case, we see that

$$u = \sum_{p \in \mathcal{P}_S} p \tag{6}$$

is the unit element of $\hat{\mathcal{G}}(S)$ (this is the image of the graph U_S), and also the unit element of the whole algebra $\hat{\mathcal{G}}$.

Claim 4.2 *Let $S \subset T$ be two finite sets. Then every $q \in \mathcal{P}_T$ resolves exactly one element of $\mathcal{P}_S$.*

Indeed, we have by (6) that

$$u = \sum_{p \in \mathcal{P}_S} p = \sum_{p \in \mathcal{P}_S} \sum_{\substack{q \in \mathcal{P}_T \\ q \text{ resolves } p}} q,$$

and also

$$u = \sum_{q \in \mathcal{P}_T} q,$$

so by the uniqueness of the representation we get that every q must resolve exactly one p.

Claim 4.3 *Let S, T, U be finite sets and $S = T \cap U$. If $x \in \hat{\mathcal{G}}(T)$ and $y \in \hat{\mathcal{G}}(U)$, then*

$$f(xy) = f(\pi_S(x)y).$$

Indeed, for every T-labeled graph G_1 and U-labeled graph G_2, the graphs $G_1 G_2$ and $\pi_S(G_1)G_2$ are isomorphic. Hence the claim follows by linearity.

Claim 4.4 *If $p \in \mathcal{P}_S$ and q resolves p, then*

$$\pi_S(q) = \frac{f(q)}{f(p)} p.$$

We show that both sides give the same inner product with every basis element in $\mathcal{P}_S$. Since q does not resolve any $p' \in \mathcal{P}_S \setminus \{p\}$, we have $p'q = 0$ for every such p'. By Claim 4.3, this implies that

$$\langle p', \pi_S(q) \rangle = f(p'\pi_S(q)) = f(p'q) = 0 = \langle p', \frac{f(q)}{f(p)} p \rangle.$$

Furthermore,

$$\langle p, \pi_S(q) \rangle = f(p\pi_S(q)) = f(pq) = f(q) = \langle p, \frac{f(q)}{f(p)} p \rangle.$$

This proves the Claim.

Claim 4.5 *Let S, T, U be finite sets and $S = T \cap U$. Then for any $p \in \mathcal{P}_S$, $q \in \mathcal{P}_{T,p}$ and $r \in \hat{\mathcal{G}}(U)$ we have*

$$f(p)f(qr) = f(q)f(pr).$$

13

Indeed, by Claims 4.4 and 4.3,

$$f(qr) = f(\pi_S(q)r) = \frac{f(q)}{f(p)} f(pr).$$

Claim 4.6 *If both $q \in \mathcal{P}_T$ and $r \in \mathcal{P}_U$ resolve p, then $qr \neq 0$.*

Indeed, by Claim 4.5,

$$f(qr) = \frac{f(q)}{f(p)} f(pr) = \frac{f(q)}{f(p)} f(r) > 0.$$

Claim 4.7 *If $S \subset T$, and $q \in \mathcal{P}_T$ resolves $p \in \mathcal{P}_S$, then $\deg(q) \geq \deg(p)$.*

It suffices to show this in the case when $|T| = |S| + 1$. Let $U \subset \mathbf{Z}_+$ be any $(|S| + 1)$-element superset of S different from T. Let Y be the set of elements in $\mathcal{P}_U$ resolving p. Then $p = \sum_{r \in Y} r$ by Claim 4.1. Here $|Y| = \deg(p)$. Furthermore, we have

$$\sum_{r \in Y} rq = q \sum_{r \in Y} r = qp = q.$$

Each of the terms on the left hand side is nonzero by Claim 4.6, and since the terms are all idempotent, each of them is the sum of one or more elements of $\mathcal{P}_{T \cup U}$. Furthermore, if $r, r' \in Y$ $(r \neq r')$, then we have the orthogonality relation

$$(rq)(r'q) = (rr')q = 0,$$

so the basic idempotents in the expansion of each term are different. So the expansion of q in $\mathcal{P}_{T \cup U}$ contains at least $|Y| = \deg(p)$ terms, which we wanted to prove.

4.2 Bounding the expansion

From now on, we assume that there is a $q > 0$ such that $r_k \leq q^k$ for all k.

So if a basic idempotent $p \in \mathcal{P}_S$ has degree D, then there are D basic idempotents on the next level with degree $\geq D$, and hence if $|T| \geq |S|$, then the dimension of $\hat{\mathcal{G}}(T)$ is at least $D^{|T \setminus S|}$. It follows that the degrees of basic idempotents are bounded by q; let D denote the maximum degree, attained by some $p \in \mathcal{P}_S$.

Let us fix such a set S and $p \in \mathcal{P}_S$ with maximum degree D. For $u \in \mathbf{Z}_+ \setminus S$, let $q_1^u, \ldots, q_D^u$ denote the elements of $\mathcal{P}_{S \cup \{u\}}$ resolving p. Note that for $u, v \in \mathbf{Z}_+ \setminus S$, there is a natural isomorphism between $\hat{\mathcal{G}}(S \cup \{u\})$ and

$\hat{\mathcal{G}}(S \cup \{v\})$ (induced by the map that fixes S and maps u onto v), and we may choose the labeling so that q_i^u corresponds to q_i^v under this isomorphism.

Next we describe, for a finite set $T \supset S$, all basic idempotents in $\mathcal{P}_T$ that resolve p. Let $V = T \setminus S$, and for every map $\phi : V \to \{1, \dots, D\}$, let

$$q_\phi = \prod_{v \in V} q_{\phi(v)}^v. \tag{7}$$

Note that by Claim 4.5,

$$f(q_\phi) = f\Big(\prod_{v \in V} q_{\phi(v)}^v\Big) = \Big(\prod_{v \in V} \frac{f(q_{\phi(v)}^v)}{f(p)}\Big) f(p) \neq 0, \tag{8}$$

and so $q_\phi \neq 0$.

Claim 4.8

$$\mathcal{P}_{T,p} = \{q_\phi : \ \phi \in \{1, \dots, D\}^V\}.$$

We prove this by induction on $|T \setminus S|$. For $|T \setminus S| = 1$ the assertion is trivial. Suppose that $|T \setminus S| > 1$. Let $u \in T \setminus S$, $U = S \cup \{u\}$ and $W = T \setminus \{u\}$; thus $U \cap W = S$. By the induction hypothesis, the basic idempotents in $\mathcal{P}_W$ resolving p are elements of the form q_ψ ($\psi \in \{1, \dots, D\}^{V \setminus \{u\}}$).

Let r be one of these. By Claim 4.6, $rq_i^u \neq 0$ for any $1 \leq i \leq D$, and clearly resolves r. We can write rq_i^u as the sum of basic idempotents in $\mathcal{P}_T$, and it is easy to see that these also resolve r. Furthermore, the basic idempotents occurring in the expression of rq_i^u and rq_j^u ($i \neq j$) are different. But r has degree D, so each rq_i^u must be a basic idempotent in $\mathcal{P}_T$ itself.

Since the sum of the basic idempotents rq_i^u ($r \in \mathcal{P}_{W,p}$, $1 \leq i \leq D$) is p, it follows that these are all the elements of $\mathcal{P}_{T,p}$. This proves the Claim.

It is immediate from the definition that an idempotent q_ϕ resolves q_i^v if and only $\phi(v) = i$. Hence it also follows that

$$q_i^v = \sum_{\phi: \ \phi(v) = i} q_\phi. \tag{9}$$

4.3 Constructing the target graph

Now we can define H as follows. Let H be the looped complete graph on $V(H) = \{1, \dots, D\}$. We have to define the node weights and edge weights.

Fix any $u \in \mathbf{Z}_+ \setminus S$. For every $i \in V(H)$, let

$$\alpha_i = \frac{f(q_i^u)}{f(p)}$$

15

be the weight of the node j. Clearly $\alpha_i > 0$.

Let $u, v \in \mathbf{Z}_+ \setminus S$, $v \neq u$, and let $W = S \cup \{u, v\}$. Let K_{uv} denote the graph on W which has only one edge connecting u and v, and let k_{uv} denote the corresponding element of $\hat{\mathcal{G}}(W)$. We can express pk_{uv} as a linear combination of elements of $\mathcal{P}_{W,p}$ (since for any $r \in \mathcal{P}_W \setminus \mathcal{P}_{W,p}$ one has $rp = 0$ and hence $rpk_{u,v} = 0$):

$$pk_{uv} = \sum_{i,j} \beta_{ij} q_i^u q_j^v.$$

This defines the weight β_{ij} of the edge ij. Note $\beta_{ij} = \beta_{ji}$ for all i, j, since $pk_{uv} = pk_{vu}$.

We prove that this weighted graph H graph gives the right homomorphism function.

Claim 4.9 *For every finite graph G, $f(G) = \hom(G, H)$.*

By (9), we have for each pair u, v of distinct elements of $V(G)$

$$pk_{uv} = \sum_{i,j \in V(H)} \beta_{i,j} q_i^u q_j^v = \sum_{i,j \in V(H)} \beta_{i,j} \sum_{\substack{\phi: \ \phi(u)=i \\ \phi(v)=j}} q_\phi = \sum_{\phi \in V(H)^V} \beta_{\phi(u),\phi(v)} q_\phi.$$

Consider now any V-labeled graph G with $V(G) = V$, and let g be the corresponding element of $\hat{\mathcal{G}}$. Then

$$pg = \prod_{uv \in E(G)} pk_{uv} = \prod_{uv \in E(G)} \Big(\sum_{\phi \in V(H)^V} \beta_{\phi(u),\phi(v)} q_\phi \Big)$$

$$= \sum_{\phi: V \to V(H)} \Big(\prod_{uv \in E(G)} \beta_{\phi(u),\phi(v)} \Big) q_\phi.$$

Since $p \in \hat{\mathcal{G}}(S)$, $g \in \hat{\mathcal{G}}(V)$ and $S \cap V = \emptyset$, we have $f(p)f(g) = f(pg)$, and so by (8),

$$f(p)f(g) = f(pg) = \sum_{\phi \in V(H)^V} \Big(\prod_{uv \in E(G)} \beta_{\phi(u),\phi(v)} \Big) f(q_\phi)$$

$$= \sum_{\phi: V \to V(H)} \Big(\prod_{uv \in E(G)} \beta_{\phi(u),\phi(v)} \Big) \Big(\prod_{v \in V(G)} \alpha_{\phi(v)} \Big) f(p),$$

The factor $f(p) > 0$ can be cancelled from both sides, completing the proof.

5 Extensions: Graphs with loops, directed graphs and hypergraphs

In our arguments we allowed parallel edges in G, but no loops. Indeed, the representation theorem is false if G can have loops: it is not hard to check that the graph parameter

$$f(G) = 2^{\#\text{loops}}$$

cannot be represented as a homomorphism function, even though its connection matrix $M(f, k)$ is positive semidefinite and has rank 1. To get a representation theorem for graphs with loops, each loop e in the target graph H must have two weights: one which is used when a non-loop edge of G is mapped onto e, and the other, when a loop of G is mapped onto e. With this modification, the proof goes through.

The constructions and results above are in fact more general; they extend to directed graphs and hypergraphs. There is a common generalization to these results, using semigroups. This will be stated and proved in a separate paper.

Acknowledgement

References

[1] M. Freedman, L. Lovász and D. Welsh (unpublished).

[2] B. Szegedy (unpublished)

[3] H. Whitney: The coloring of graphs, *Ann. of Math.* **33** (1932), 688-718.